

# Accepted Manuscript

Bayesian updating rules and AGM belief revision

Pathikrit Basu

PII: S0022-0531(18)30680-X  
DOI: <https://doi.org/10.1016/j.jet.2018.11.005>  
Reference: YJETH 4847

To appear in: *Journal of Economic Theory*

Received date: 25 February 2018  
Revised date: 19 November 2018  
Accepted date: 22 November 2018

Please cite this article in press as: Basu, P. Bayesian updating rules and AGM belief revision. *J. Econ. Theory* (2018), <https://doi.org/10.1016/j.jet.2018.11.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# Bayesian updating rules and AGM belief revision\*

Pathikrit Basu<sup>†</sup>

November 26, 2018

## Abstract

We interpret the problem of updating beliefs as a choice problem (selecting a posterior from a set of admissible posteriors) with a reference point (prior). We use AGM belief revision to define the support of admissible posteriors after arrival of information, which applies also to zero probability events. We study two classes of updating rules for probabilities : 1) "lexicographic" updating rules where posteriors are given by a lexicographic probability system 2) "minimum distance" updating rules which select the posterior closest to the prior by some metric. We show that an updating rule is lexicographic if and only if it is Bayesian, AGM-consistent and satisfies a weak form of path independence. While not all lexicographic updating rules have a minimum distance representation, we study a sub-class of lexicographic rules, which we call "support-dependent" rules, which admit a minimum distance representation. Finally, we apply our approach to the problem of updating preferences.

Keywords : Bayesian Updating, AGM Belief Revision, Information Processing, Lexicographic Updating

JEL : D01, D80, D81, D83

## 1 Introduction

Updating beliefs in light of newly acquired information is a problem that is relevant and occurs in many situations in economic theory. In most environments, an agent's belief is

---

\*I am deeply grateful to Kalyan Chatterjee, Edward Green, Neil Wallace, Stephen Simpson, Giacomo Bonnano, Ariel Rubinstein, Arunava Sen, Vijay Krishna, Debasis Mishra, Federico Echenique, Yuval Salant, Willemien Kets, Ryan Fang and Bruno Salcedo for discussions, comments and suggestions.

<sup>†</sup>California Institute of Technology. E-mail : pathkrtb@caltech.edu. Mail Address : 202, Annenberg Center for Information Science and Technology, 330 S Chester Ave, Pasadena, CA 91125.

represented by a probability measure over a state space, elements of which are payoff relevant. As a result, actions of the agent depend crucially on his opinion or belief over the state variable and consequently also on how he chooses to update his belief upon learning an event. A dominant principle used to update probabilities in most models is Bayesian updating. Starting with a prior  $\pi$  and observing a positive probability event  $A$ , Bayesian updating suggests the posterior  $\pi'(E) = \pi(E \cap A)/\pi(A)$ . However, in the event of a surprise i.e. observing a zero probability event, Bayesian updating remains silent and is not well-defined. Such situations arise in extensive-form games of imperfect information where a player's strategy must specify which action to choose in an information set that is reached with zero probability and the updating problem is one of assigning probabilities to nodes in the information set. The solution concepts of sequential equilibrium and trembling-hand perfect equilibrium (see Kreps and Wilson [1982b] and Selten [1975]) both place restrictions on admissible beliefs on information sets which lie off path (using Bayesian updating otherwise). These put restrictions on off-path beliefs by requiring them to be the limit of on-path beliefs corresponding to a sequence of perturbations of the equilibrium strategy profile. In perfect information extensive-form games as well, due to strategic uncertainty, a player may revise their beliefs after observing a deviation, about their opponent's strategies and hence, their future actions (see Battigalli and Siniscalchi [2002]). Abstracting away from the game-theoretic scenario, we ask whether in the probabilistic model itself, there exists a systematic way to extend Bayesian updating to zero probability events.

In this paper, we interpret the updating problem as a choice problem with a reference point (see Rubinstein and Zhou [1999])). The reference point here is a prior  $\pi$  on the state space and given an event  $E$ , the choice problem is one of choosing a posterior from an admissible set of posteriors which have a common support in  $E$ . This raises the question of the how the admissible set of posteriors should be selected for a given updating problem  $(\pi, E)$ . Since we wish to *extend* Bayesian updating, if  $\pi(E) > 0$  we would want this set to be all probability measures whose support is the intersection of the support of the prior and the observed positive probability event and choice would be the Bayesian posterior. The intersection would include all states consistent both with the information  $E$  and the prior belief  $\pi$ . But what happens if  $E$  has zero probability? Is there a consistent way to select the set of admissible posteriors for all events  $E$ ? This would allow us to define the support of the posterior belief after both positive and zero probability events.

The question of finding the support of the posterior can be posed as a problem of "theory change" as is known in the non-probabilistic theory of belief revision from propositional

logic, namely, AGM belief revision (see Alchourron et al. [1985]). In that setting, an agent's primitive is a belief set, which is a set of propositions or events that the agent believes to be true (in the present context this would be all probability one events according to the prior). The belief revision problem is one of revising the belief set to a new belief set, based on the information that event  $E$  has occurred. The desirable feature of applying the AGM procedure is that it provides solutions to the belief revision problem even when information obtained is inconsistent with prior beliefs. This property of the theory is, as we shall see, intimately linked with the problem of updating over zero probability events. If the prior is  $\pi$  and  $\mathcal{K}(\pi)$  is its associated belief set, then we want the posterior belief  $\pi'$  to be such that the now revised belief set,  $\mathcal{K}(\pi')$ , is derived from  $\mathcal{K}(\pi)$  and  $E$  via AGM belief revision.

The consistency requirement here means that we want the solution to our probabilistic updating problem to be consistent with the underlying change in the associated belief sets. If we believe that this latter change should satisfy certain postulates and properties (which is a central concern of the literature on belief revision, see Costa and Pedersen [2011]), then we obtain a set of restrictions on the posteriors to be considered admissible. The advantage of this approach is that it is well defined in specifying the support of posteriors after all events, including zero-probability events. It also has the desirable property that all admissible posteriors have common support in the event observed and hence, updating satisfies consequentialism. When an updating rule abides by the AGM procedure, we say that it is *AGM-consistent*.

We focus on two classes of updating rules. The first class of updating rules we study are *lexicographic* updating rules where the posteriors are defined by a lexicographic probability system (see Blume et al. [1991a], Halpern [2010] and Hammond [1994]). We show that an updating rule is lexicographic if and only if it is Bayesian, AGM-consistent and satisfies a weak form of path independence (order in which information arrives does not matter). The weakening is in the sense that order independence is satisfied only for certain pairs of events. It also turns out that this weakening is crucial. There exist no updating rules which are Bayesian and satisfy strong path independence (order-independence on all consistent pairs of events). Extending this approach to the problem of updating preferences, we also obtain a similar characterisation of lexicographic updating in that setting.

The second class of updating rules, which we call *minimum-distance*, picks the posterior closest to the prior according to a metric defined on the space of probability measures<sup>1</sup>. We

---

<sup>1</sup>Similar rules have also been studied by Perea [2009] and Rubinstein and Zhou [1999].

study the relationship between lexicographic and minimum distance updating rules. While not all lexicographic updating rules have a minimum distance representation, we show that a sub-class of lexicographic rules, which we call *support-dependent* updating rules, indeed admit minimum distance representations. In these rules, the alternative hypotheses that are used to update over events of zero probability, depend only on the support of the prior. Hence, they depend only on the set of states which have ex-ante zero probability.

For an exposition of AGM theory, Costa and Pedersen [2011]) and Huber [2013]) provide excellent surveys. The problem of choosing supports of posteriors in a manner consistent with AGM belief revision has been studied by Bonanno [2009]), where a choice correspondence given an event, chooses as a subset of the event, the support of the posterior. The framework is non-probabilistic and it is shown that rationalisability of the choice correspondence is equivalent to AGM-consistency. In the present framework, we derive a counterpart of this result and it provides a very useful characterisation of AGM-consistency of an updating rule (see also Grove [1988]). The relationship between lexicographic probability systems and AGM consistency has also been discussed in Shoham and Leyton-Brown [2009]) in terms of the belief operator for revising belief sets, again, in a non-probabilistic framework. We derive and build on that observation in our framework and also establish a complete characterisation of lexicographic updating rules.

In the theory of decision under uncertainty, in addition to Blume et al. [1991a]), who consider lexicographic probabilities, there has been some attention devoted to alternative approaches to dealing with zero probability events. Myerson [1986]) provides axiomatic foundations for conditional probability systems. Ortoleva [2012] studies an alternative approach where once a zero probability occurs, the agent uses a belief over beliefs and the maximum likelihood rule to obtain posteriors. In the present work, such updating rules may violate AGM consistency, which is central to the analysis considered here. Karni and Vierø [2013] study a framework where the set of states itself can expand due to growing awareness and they consider the phenomenon of reverse Bayesianism. This requires that whenever the state space grows and the support of the prior belief is contained in that of the posterior, the prior can be obtained by applying Bayes rules to the posterior. The issue of updating ambiguous beliefs as defined in Schmeidler [1989] and Gilboa and Schmeidler [1989] has been discussed in Gilboa and Schmeidler [1993]. They consider the problem of updating convex non-additive probabilities and establish the equivalence of the dempster-shafer rule for conditioning and maximum likelihood updating. Though in the present work we do not discuss ambiguous beliefs, our approach may be used to define updating rules for it. This extension

could potentially provide us with a connection between ambiguity aversion and agents' attitudes to zero probability events. Finally, our extension involving updating preferences is also related to the above papers and Hanany and Klibanoff [2007], who study updating rules for preferences with multiple priors.

The outline of this paper is as follows : In section 2, we introduce the framework and provide a brief summary of AGM belief revision. In section 3, we study lexicographic updating rules and minimum distance updating rules are studied in Section 4. Finally, we study preference updating in Section 5. Some proofs are in the appendix.

## 2 Framework

### 2.1 Updating Rules

In this section, we describe the environment and introduce the updating problem faced by the agent. We denote as  $\Omega$ , a non-empty finite set, to be the underlying state space on which the agent has beliefs. An agent's probabilistic belief about the state space is defined by a probability measure  $\pi \in \Delta(\Omega)$ . We shall sometimes call  $\pi$ , the agent's *prior* to emphasise that  $\pi$  is his belief prior to arrival of information. Similarly, we shall refer to a belief  $\pi'$  as the agent's *posterior* when it is formed after information arrives. Throughout the paper, we shall denote as  $\text{supp}(\pi) := \{\omega \in \Omega : \pi(\omega) > 0\}$ , the support of a probability measure  $\pi$ . Upon learning that an event  $A \subseteq \Omega$  has taken place, the agent updates his prior belief  $\pi$  over the state space  $\Omega$  to a posterior belief  $\pi' \in \Delta(\Omega)$ . The agent performs this task in two stages. Upon observing  $A$ , he lists a *menu* of possible posteriors  $\Pi(\pi, A) \subseteq \Delta(\Omega)$  to choose from. From the menu, he selects a single posterior  $\pi' \in \Pi(\pi, A)$ . This two-stage choice completely describes his belief updating procedure and we define it formally below.

**Definition 2.1.** An *updating rule* is a tuple  $\langle c, \Pi \rangle$  where  $c : \Delta(\Omega) \times 2^\Omega \setminus \{\emptyset\} \rightarrow \Delta(\Omega)$  and  $\Pi : \Delta(\Omega) \times 2^\Omega \rightarrow 2^{\Delta(\Omega)}$  such that :

- For all  $\pi', \pi'' \in \Pi(\pi, A) : \text{supp}(\pi') = \text{supp}(\pi'') \subseteq A$ .
- $c(\pi, A) \in \Pi(\pi, A)$ .

The first condition says that the menu of posteriors selected by the agent has a common support that is contained completely in the event learned  $A$ . The common support assumption stems from the rationale that the set of states the agent cannot disregard must be independent of the choice of the posterior and depends solely on  $(\pi, A)$ , otherwise additional

information must have influenced him to do so. Inclusion of the support in  $A$  captures the idea that the event  $A$  has been observed to have taken place and the agent disregards as improbable, states outside it. This latter condition may be viewed as updating satisfying consequentialism. In what follows, we shall sometimes write  $\text{supp}\Pi(\pi, A)$  to denote the common support of the probability measures in  $\Pi(\pi, A)$ .

Note that an updating rule has two components :  $c$  and  $\Pi$ . One may view  $\Pi$  as capturing the "qualitative" aspect of the updating process. This corresponds to identifying the support of the posterior distribution. The function  $c$  captures the "quantitative" part, in the sense that  $c(\pi, A)$  tells us the precise numerical values for the posterior probabilities supported on  $\text{supp}\Pi(\pi, A)$ . This decomposition of the updating process seems natural particularly when dealing with zero probability events. An agent may update beliefs by first determining which states to place positive weight on and then choosing the exact distribution of probabilities. For example, in the context of games, an updating rule may be applied as follows. After a deviation occurs,  $\Pi$  tells us a player's revised belief about the set of strategies that the opponent could be playing with positive probability. Then,  $c$  assigns weights to this set of probable strategies.

## 2.2 Bayesian Updating

We now discuss updating rules defined by Bayesian updating. Such updating rules apply Bayes' rule when positive probability events are observed. The support of the posterior then equals the intersection of the support of the prior and the event observed.

**Definition 2.2.** An updating rule  $\langle c, \Pi \rangle$  is said to be *Bayesian* if :

- Whenever  $\pi(A) > 0$ ,  $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\pi) \cap A\}$ .
- Moreover, when  $\pi(A) > 0$ ,  $c(\pi, A)(\omega) = \frac{\pi(\{\omega\} \cap A)}{\pi(A)}$ .

Notice in the above definition, no restriction has been placed on how the updating rule behaves on zero probability events. Additionally, in this case, we make no restrictions on the common support of the menu of posteriors chosen by the individual. This makes the class of Bayesian updating rules very large and potentially allows for varied attitudes in updating when agents are completely surprised. Restrictions on  $c(\pi, A)$  and  $\Pi(\pi, A)$  when  $\pi(A) = 0$  can potentially have very strong implications and is indeed the main subject of the remainder of the paper. We impose *consistency* requirements on updating rules to restrict behaviour on zero probability events. One would hope for such additional requirements to not be dictated by arbitrary exogenous rules but arise from the primitives  $(\pi, A)$  in a



manner that is plausible. In doing so, we impose that the updating rule be consistent with a belief revision procedure proposed by Alchourron et al. [1985] (henceforth AGM), that widely arises in propositional logic in the context of updating sets of propositions ("belief sets" in the parlance of the belief revision literature). We introduce a notion called *AGM-consistency* for updating rules based on this requirement. The next section is devoted to its formulation and interpretation in the present environment.

## 2.3 AGM Consistency

In this section, we introduce a notion of consistency of updating rules based on AGM belief revision. We will first provide some definitions and then describe the revision procedure. It shall be useful to interpret subsets of the state space  $\Omega$  as an algebra of propositions with an event  $E$  being treated as a proposition. The set of all propositions is thus  $2^\Omega$ . In this context, the notion of logical consequence may be defined as follows :

**Definition 2.3.** Let  $\mathcal{G} \subseteq 2^\Omega$  and let  $A \subseteq \Omega$ . We say  $\mathcal{G} \vdash A$  (meaning  $A$  is a logical consequence of  $\mathcal{G}$ ) if  $\bigcap_{G \in \mathcal{G}} G \subseteq A$  and define the consequence operator as  $Cn(\mathcal{G}) = \{A : \mathcal{G} \vdash A\}$ .

We say a set of events  $\mathcal{G}$  is *logically closed* if  $Cn(\mathcal{G}) \subseteq \mathcal{G}$  and *consistent* if  $\bigcap_{G \in \mathcal{G}} G \neq \emptyset$ .

Let us consider the following example.

**Example 2.1.** Suppose  $\Omega = \{x, y, z\}$ . Then, the collection  $\mathcal{G} = \{\{x, y\}, \{x, y, z\}\}$  is logically closed. Indeed,  $\bigcap_{E \in \mathcal{G}} E = \{x, y\}$  and  $\{x, y\}$  and  $\{x, y, z\}$  are the only sets which contain the intersection. Now, consider  $\mathcal{G}' = \{\{x\}, \{x, y\}, \{x, y, z\}\}$ . This is not logically closed. The set  $\{x, z\}$  is not in  $\mathcal{G}'$  however  $\bigcap_{E \in \mathcal{G}'} E = \{x\} \subseteq \{x, z\}$ . Note also that both  $\mathcal{G}$  and  $\mathcal{G}'$  are consistent.

Let us discuss the notion of logical consequence defined above. A set  $A$  is said to be a logical consequence of  $\mathcal{G}$  if it is impossible that all events in  $\mathcal{G}$  occur but  $A$  does not occur i.e. there does not exist any state  $\omega$  such that  $\omega \in \bigcap_{G \in \mathcal{G}} G$  but  $\omega \notin A$ . This is a definition which is natural and analogous to the definition of logical consequence in propositional logic if we

---

<sup>2</sup>Consider, for example, tossing a coin one hundred times. The state space would be the set  $\{H, T\}^{100}$ . Now, the propositions "the 24th coin toss led to Heads" and "the 26th coin toss led to tails" can both be expressed as events  $\{\omega | \omega_{24} = H\}$  and  $\{\omega | \omega_{26} = T\}$ . The conjunction of the two propositions would correspond to the intersection of the two events.



were to interpret the set of all truth valuations as the state space  $\Omega^3$ . In the remainder of the paper, a logically closed and consistent set of events will often be referred to as a *theory* or *belief set* and will be denoted as  $\mathcal{K}$ . Corresponding to each  $\pi \in \Delta(\Omega)$ , we can define a belief set consisting of all the probability one events.

**Definition 2.4.** Let  $\Omega$  be finite and let  $\pi \in \Delta(\Omega)$ . The *belief set corresponding to  $\pi$*  is defined as the set  $\mathcal{K}(\pi) = \{E \subseteq \Omega : \pi(E) = 1\}$

It follows that  $\mathcal{K}(\pi)$  is logically closed and consistent. Clearly,  $\bigcap_{E \in \mathcal{K}(\pi)} E = \text{supp}(\pi)$ . The set  $\mathcal{K}(\pi)$  is logically closed since all probability one events contain the support  $\text{supp}(\pi)$ . In fact,  $\mathcal{K}(\pi)$  contains exactly the probability one events. Consistency of  $\mathcal{K}(\pi)$  follows from the fact that the support is non-empty. Consider the example below.

**Example 2.2.** Suppose  $\Omega = \{x, y, z\}$  and  $\pi = (\lambda, 1 - \lambda, 0) \in \Delta(\Omega)$  where  $\lambda \in (0, 1)$ . Then,  $\mathcal{K}(\pi) = \{\{x, y\}, \{x, y, z\}\}$  and from Example 2.1, it follows that  $\mathcal{K}(\pi)$  is logically closed and consistent.

### 2.3.1 AGM Postulates

The theory of belief revision (see Costa and Pedersen [2011] and Huber [2013] for introductory surveys) is devoted to studying the revision of belief sets  $\mathcal{K}$  upon learning a new event  $A$  by means of *revision operator*  $*$ . Formally, letting  $\mathcal{T}$  denote the set of all logically closed and consistent belief sets, a revision operator is a function  $* : \mathcal{T} \times 2^\Omega \setminus \{\emptyset\} \rightarrow 2^{2^\Omega}$ . Hence, revision of the belief set  $\mathcal{K}$  leads to the new belief set  $\mathcal{K} * A$ . AGM study revision operators  $*$  that satisfy the following postulates :

1. (*Closure*)  $\mathcal{K} * A = \text{Cn}(\mathcal{K} * A)$ .
2. (*Success*)  $A \in \mathcal{K} * A$ .
3. (*Inclusion*)  $\mathcal{K} * A \subseteq \text{Cn}(\mathcal{K} \cup \{A\})$ .
4. (*Vacuity*) If  $A^c \notin \mathcal{K}$ , then  $\text{Cn}(\mathcal{K} \cup \{A\}) \subseteq \mathcal{K} * A$ .
5. (*Consistency*) If  $A \neq \emptyset$ , then  $\mathcal{K} * A$  is consistent.

---

<sup>3</sup>See, for example, Chapter 2 of Gallier [2015]. For a set of atomic propositions,  $\mathbb{P}$ , the state space would be all truth valuations  $\{T, F\}^\mathbb{P}$ . Each proposition  $\alpha$  derived from  $\mathbb{P}$  through elementary operations of conjunction, disjunction and negation, has an event corresponding to it. Namely, it is the event  $E(\alpha) = \{\omega | \omega(\alpha) = T\}$ . The converse is true as well. For any event  $E$ , there exists a proposition  $\alpha$ , such that  $E = E(\alpha)$ .

6. (*Extensionality*) If  $Cn(\{A\}) = Cn(\{B\})$ , then  $\mathcal{K} * A = \mathcal{K} * B$ .
7. (*Superexpansion*)  $\mathcal{K} * (A \cap B) \subseteq Cn((\mathcal{K} * A) \cup \{B\})$ .
8. (*Subexpansion*) If  $B^c \notin \mathcal{K} * A$ , then  $Cn((\mathcal{K} * A) \cup \{B\}) \subseteq \mathcal{K} * (A \cap B)$ .

The above axioms on the revision procedure have come to be known in the literature as the *AGM postulates*. The Closure postulate says that the revised belief set  $\mathcal{K} * A$  is logically closed. Success guarantees that the information  $A$  is contained in  $\mathcal{K} * A$ . Inclusion means that any event in the belief set  $\mathcal{K} * A$  is implied by  $\mathcal{K}$  and  $A$  together. Vacuity means that if the event  $A$  is consistent with  $\mathcal{K}$ , then the revised set  $\mathcal{K} * A$  contains all the consequences of  $\mathcal{K}$  and  $A$ . Consistency requires that  $\mathcal{K} * A$  is consistent. In the current context with events, Extensionality is trivially satisfied since  $Cn(\{A\}) = Cn(\{B\})$  implies  $A = B$ . Finally, note that the postulates Superexpansion and Subexpansion are similar to the postulates Inclusion and Vacuity.

In their seminal paper, AGM (Alchourron et al. [1985]), establish that adherence to the postulates is equivalent to following a two stage procedure for belief revision involving a *contraction* stage and *expansion* stage. The contraction step deletes events from  $\mathcal{K}$  by selecting from maximal subsets (according to set inclusion) of  $\mathcal{K}$  consistent with  $A$  via a selection function, and treats the intersection of the selected maximal subsets as the set of retained events. The expansion stage adds  $A$ , by union, to the retained set of events from the contraction stage and treats the logical closure of the resultant set as the revised belief set  $\mathcal{K} * A$ . For further details, see, for example, Costa and Pedersen [2011].

### 2.3.2 Incorporating AGM as a consistency requirement

We now define what it means for an updating rule to be AGM-consistent.

**Definition 2.5.** An updating rule  $\langle c, \Pi \rangle$  is said to be *AGM-consistent* if there exist revision operators  $\{*_\pi\}_{\pi \in \Delta(\Omega)}$  (one for each prior  $\pi$ ) satisfying the AGM postulates, such that for all  $(\pi, A)$ , we have that

$$\Pi(\pi, A) = \{\pi' : \mathcal{K}(\pi') = \mathcal{K}(\pi) *_\pi A\}. \quad (1)$$

The requirement is that the choice of posterior must be such that its support is consistent with AGM belief revision i.e the posterior  $\pi'$  should be such that the belief set corresponding to it equals the belief set obtained by revising the belief set corresponding to the prior  $\pi$  based on event  $A$ . Figure 1 depicts the requirement by means of a diagram. AGM consistency

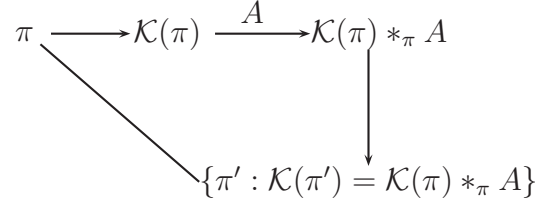


Figure 1: AGM-Consistency

requires the use of revision operators, which as mathematical objects, could be potentially be very difficult to handle. The following result, which follows from a theorem due to Grove (see Theorems 1 and 2 in Grove [1988]), known as Grove's representation theorem, allows a simple characterisation of AGM-consistent updating rules involving complete and transitive relations on the state space  $\Omega$ .

**Proposition 1.** (Grove) *An updating rule  $\langle c, \Pi \rangle$  is AGM-consistent if and only if there exists a family  $\{\geq_\pi\}_{\pi \in \Delta(\Omega)}$  of complete and transitive relations on  $\Omega$  such that for all  $\pi$ , we have  $M_{\geq_\pi}(\Omega) = \text{supp}(\pi)$  and*

$$\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = M_{\geq_\pi}(A)\}, \quad (2)$$

where  $M_{\geq_\pi}(A) = \{\omega \in A : \omega \geq_\pi \omega' \text{ for all } \omega' \in A\}$  is the set of  $\geq_\pi$ -greatest elements in  $A$ .

The diagram in Figure 2 explains the idea behind the above representation. The concentric bands around  $\text{supp}(\pi)$  represent the equivalence classes of  $\geq_\pi$ , in decreasing order, from the inner to outer bands. When  $A$  is observed, we find the highest equivalence class that intersects with  $A$ . This intersection is equal to  $\text{supp}(\pi, A)$ .

Finally, note that AGM consistency only places a restriction on the support of the posterior. Moreover, the support depends on the sets  $\mathcal{K}(\pi)$ ,  $A$  and the ordering  $\geq_\pi$ . Here, note that  $\mathcal{K}(\pi)$  only depends on the support of  $\pi$ . However, apart from  $\text{supp}(\pi)$ , the equivalence classes of  $\geq_\pi$  could potentially depend on the distribution of  $\pi$ . Hence, reactions to zero probability events could potentially depend on the distribution of the prior.

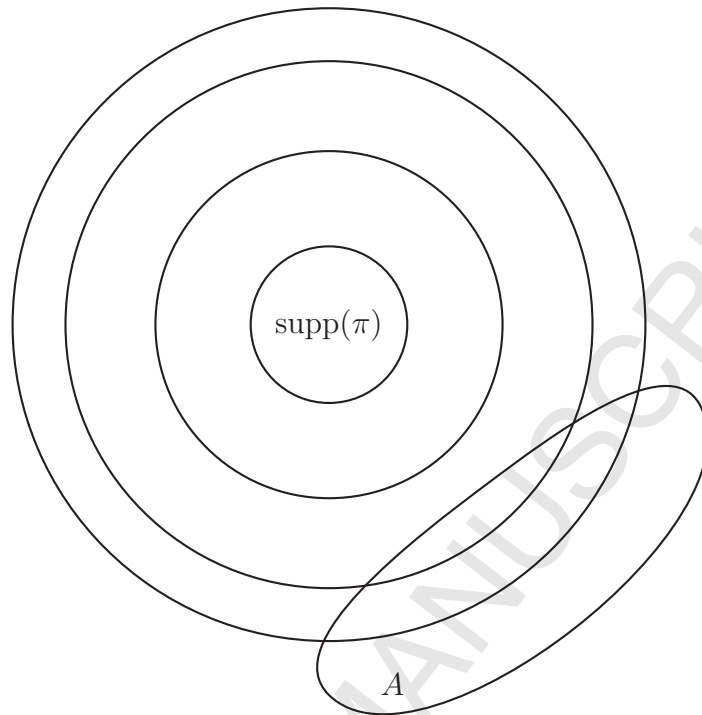


Figure 2: AGM consistency in terms of the ordering  $\geq_\pi$

### 3 Lexicographic Updating Rules

In this section, we study updating rules based on updating using lexicographic probability systems. The main result is that an updating rule is lexicographic if and only if it is Bayesian, AGM consistent and satisfies a weak form of path independence i.e the order in which information arrives does not influence the final posterior. Path independence is a key property in the characterisation. We first show that when a stronger version is imposed, there exist no Bayesian updating rules which satisfy path independence (Claim 1). Propositions 2 and 3 correspond to the main result of this section.

#### 3.1 Path Independence

A desirable property for an updating rule to have is path independence. This means that the order in which information is received does not affect the final posterior. On positive probability events, indeed Bayesian updating rules satisfy this property of order independence. We require that it be satisfied for all consistent pairs of events, including zero probability events according to the prior. We define this formally.

**Definition 3.1.** An updating rule  $\langle c, \Pi \rangle$  satisfies *strong path independence* if for all

$A, B \subseteq \Omega$  such that  $A \cap B \neq \emptyset$  and for all  $\pi \in \Delta(\Omega)$  it is true that

$$c(c(\pi, A), B) = c(c(\pi, B), A).$$

We obtain the following result.

**Claim 1.** *If  $|\Omega| \geq 3$ , then there exists no Bayesian updating rule which satisfies strong path independence.*

*Proof.* For any states  $\omega \in \Omega$ , let  $\delta_\omega$  denote the Dirac-measure, which places probability one on  $\omega$ . Let  $\langle c, \Pi \rangle$  be a Bayesian updating rule. Consider three distinct states  $\{\omega_1, \omega_2, \omega_3\} \subseteq \Omega$  and consider the prior  $\delta_{\omega_1}$ . Now define the sets  $A = \{\omega_1, \omega_2\}$ ,  $B = \{\omega_1, \omega_3\}$ ,  $C = \{\omega_2, \omega_3\}$ . Since  $\langle c, \Pi \rangle$  is Bayesian, it is true that  $c(\delta_{\omega_1}, A) = c(\delta_{\omega_1}, B) = \delta_{\omega_1}$ . Now define  $\hat{\pi} = c(\delta_{\omega_1}, C) = c(c(\delta_{\omega_1}, A), C) = c(c(\delta_{\omega_1}, B), C)$ . Now by definition it is the case that  $\text{supp}(\hat{\pi}) \subseteq C = \{\omega_2, \omega_3\}$ . There are now three cases.

- *Case 1 :*  $\text{supp}(\hat{\pi}) = C = \{\omega_2, \omega_3\}$ . Now since  $\langle c, \Pi \rangle$  is Bayesian,  $c(c(\delta_{\omega_1}, C), B) = \delta_{\omega_3}$ . But, we have  $c(c(\delta_{\omega_1}, B), C) = c(\delta_{\omega_1}, C) = \hat{\pi} \neq \delta_{\omega_3}$ .
- *Case 2 :*  $\text{supp}(\hat{\pi}) = \{\omega_2\}$ . This implies that  $\hat{\pi} = \delta_{\omega_2}$ . Now,  $c(c(\delta_{\omega_1}, C), B) = c(\delta_{\omega_2}, B)$ . Note that  $\text{supp}(c(\delta_{\omega_2}, B)) \subseteq B$ . So clearly  $c(\delta_{\omega_2}, B) \neq \delta_{\omega_2}$ . But we have  $c(c(\delta_{\omega_1}, B), C) = c(\delta_{\omega_1}, C) = \delta_{\omega_2}$ .
- *Case 3 :*  $\text{supp}(\hat{\pi}) = \{\omega_3\}$ . This implies that  $\hat{\pi} = \delta_{\omega_3}$ . Now,  $c(c(\delta_{\omega_1}, C), A) = c(\delta_{\omega_3}, A)$ . Note that  $\text{supp}(c(\delta_{\omega_3}, A)) \subseteq A$ . So clearly  $c(\delta_{\omega_3}, A) \neq \delta_{\omega_3}$ . But we have  $c(c(\delta_{\omega_1}, A), C) = c(\delta_{\omega_1}, C) = \delta_{\omega_3}$ .

□

The above result may be viewed as unsettling. Note that when restricted to positive probability events, Bayesian updating satisfies path independence. However, when we expand the domain of updating to allow for zero probability events, path independence cannot be satisfied. Note also that with two states of the world, the result does not apply. Moreover, with two states, one can verify that every Bayesian updating rule satisfies strong path independence.

### 3.2 Lexicographic Updating Rules : Characterisation

In Blume et al. [1991a], updating via lexicographic probability systems (LPS) is defined and axiomatic foundations are studied for a model of decision making under uncertainty where agents rank acts according to lexicographic expected utility. LPS's provide a new class of

updating rules where the posterior  $c(\pi, A)$  is provided by the Bayesian update of the first probability measure in a hierarchy  $\langle p_1, \dots, p_K \rangle$  with  $p_1 = \pi$ , where  $A$  assumes a positive probability i.e.  $p_k(A) > 0$ . We now investigate the relationship between LPS's and AGM-consistent updating rules. First, we define formally what it means for an updating rule to be lexicographic.

**Definition 3.2.** A *lexicographic probability system* (LPS) is a finite sequence  $\mathbf{p} = \langle p_1, \dots, p_K \rangle$  of probability measures on  $\Omega$  such that for any non-empty event  $A \subseteq \Omega$ , there exists an  $1 \leq i \leq K$  such that  $p_i(A) > 0$ . A *lexicographic conditional probability system* (LCPS) is an LPS where the supports of the probability measures in  $\langle p_1, \dots, p_K \rangle$  are pairwise disjoint.

LPS's can be used to define posteriors for all non-empty events in the following way. For  $A \subseteq \Omega$  non-empty, define the posterior to be  $\mathbf{p}(\cdot|A) := p_{k^*}(\cdot|A)$  where  $k^* = \min\{k : p_k(A) > 0\}$ .

**Claim 2.** For every LPS  $\mathbf{p}$  there exists an LCPS  $\mathbf{q}$  such that for all non-empty  $A \subseteq \Omega$ ,  $\mathbf{p}(\cdot|A) = \mathbf{q}(\cdot|A)$ .

*Proof.* Let  $\mathbf{p} = \langle p_1, \dots, p_K \rangle$  be an LPS. Define  $\mathbf{q}$  as follows :

- Set  $q_1 = p_1$ .
- If  $\Omega \setminus \cup_{i=1}^k \text{supp}(q_i) \neq \emptyset$ , then define  $q_{k+1} = p_l(\cdot | \text{supp}(p_l) \setminus \cup_{i=1}^k \text{supp}(q_i))$  where  $l = \min\{j : \text{supp}(p_j) \setminus \cup_{i=1}^k \text{supp}(q_i) \neq \emptyset\}$ . If  $\Omega \setminus \text{supp}(q_k) = \emptyset$ , then stop.

Since  $\Omega$  is finite, this process terminates at some step  $K$ . Define  $\mathbf{q} = \langle q_1, \dots, q_K \rangle$ . It can now be shown that indeed  $\mathbf{p}(\cdot|A) = \mathbf{q}(\cdot|A)$  for all non-empty  $A \subseteq \Omega$   $\square$

It can also be shown that if two LCPS's  $q$  and  $q'$  induce the same conditional probabilities, then  $q = q'$ . The implication of this is that for any LPS  $p$  there exists a unique LCPS  $q$  which induces the conditional probabilities as  $p$ .

**Definition 3.3.** An updating rule  $\langle c, \Pi \rangle$  is said to be *lexicographic* if there exists a family  $\{\mathbf{p}^\pi\}_{\pi \in \Delta(\Omega)}$  of LPS's such that :

- $p_1^\pi = \pi$  for all  $\pi \in \Delta(\Omega)$ .
- $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\mathbf{p}^\pi(\cdot|A))\}$ .
- $c(\pi, A) = \mathbf{p}^\pi(\cdot|A)$ .

Let us now make some useful observations. Firstly, note that any family of LPS's  $\{\mathbf{p}^\pi\}_{\pi \in \Delta(\Omega)}$  yields an updating rule defined by the conditions above. Secondly, suppose for two families  $\{\mathbf{p}^\pi\}_{\pi \in \Delta(\Omega)}$ ,  $\{\mathbf{q}^\pi\}_{\pi \in \Delta(\Omega)}$  we have that  $\mathbf{p}^\pi(\cdot|A) = \mathbf{q}^\pi(\cdot|A)$  for all non-empty  $A \subseteq \Omega$  and for all  $\pi \in \Delta(\Omega)$ . Then, both families induce the same updating rule. Having made these observations, we are now ready to prove our first result.

**Proposition 2.** *Every lexicographic updating rule  $\langle c, \Pi \rangle$  is Bayesian and AGM-consistent.*

*Proof.* Let  $\langle c, \Pi \rangle$  be lexicographic with the corresponding family  $\{\mathbf{p}^\pi\}_{\pi \in \Delta(\Omega)}$ . From the above claim we can construct a family of LCPS's  $\{\mathbf{q}^\pi\}_{\pi \in \Delta(\Omega)}$  such that  $\mathbf{p}^\pi(\cdot|A) = \mathbf{q}^\pi(\cdot|A)$  for all non-empty  $A \subseteq \Omega$  and for all  $\pi \in \Delta$ . Hence  $\{\mathbf{q}^\pi\}_{\pi \in \Delta(\Omega)}$  the same updating rule.

Let  $\pi$  and  $A \subseteq \Omega$  be such that  $\pi(A) > 0$ . Now  $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\pi) \cap A\}$  and  $c(\pi, A) = p_1^\pi(\cdot|A) = q_1^\pi(\cdot|A) = \pi(\cdot|A)$ . So  $\langle c, \Pi \rangle$  is Bayesian. We now show that it is AGM-consistent. For  $\pi \in \Delta(\Omega)$ , define the relation

$$\omega \geq_\pi \omega' \iff \omega \in \text{supp}(q_k^\pi) \text{ and } \omega' \in \text{supp}(q_l^\pi) \text{ for } k \leq l.$$

Since the collection  $\{\text{supp}(q_k^\pi)\}_k$  constitutes a partition of  $\Omega$ , the order  $\geq_\pi$  is complete and transitive. Denote as  $M_{\geq_\pi}^k(\Omega)$ , the set of all  $k$ -th highest states according to  $\geq_\pi$ . Clearly,  $M_{\geq_\pi}^k(\Omega) = \text{supp}(q_k^\pi)$  for all  $k$ . Let  $k(\pi, A) = \min\{k : q_k^\pi(A) > 0\}$ . From this observation we get that  $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(\mathbf{q}^\pi(\cdot|A))\} = \{\pi' : \text{supp}(\pi') = M_{\geq_\pi}^{k(\pi, A)}(\Omega) \cap A = M_{\geq_\pi}(A)\}$ . Hence, from Proposition 1, it follows that  $\langle c, \Pi \rangle$  is AGM-consistent.  $\square$

We have established that every lexicographic updating rule is Bayesian and AGM-consistent. However, there exist Bayesian and AGM-consistent updating rules which are not lexicographic (see Example 3.1). One natural question to address at this point would be : under what additional conditions are Bayesian, AGM-consistent updating rules lexicographic? It turns out that the only additional condition we need is a weak form of path independence. We first define path independence and then derive our result.

**Definition 3.4.** An updating rule  $\langle c, \Pi \rangle$  satisfies *weak path independence* if for all  $\pi \in \Delta(\Omega)$  and for all  $A, B \subseteq \Omega$  such that  $\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset$ , it is true that

$$c(c(\pi, A), B) = c(c(\pi, B), A).$$

Some remarks about weak path independence are in order. Note by definition, we have  $\text{supp}\Pi(\pi, A) \subseteq A$  and  $\text{supp}\Pi(\pi, B) \subseteq B$ . Hence, indeed weak path independence is a weaker notion compared to strong path independence. Moreover, if an updating rule is AGM consistent, then it is true that for all  $\pi \in \Delta(\Omega)$  and for all  $A, B \subseteq \Omega$  such that  $\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset$ ,  $\text{supp}(c(c(\pi, A), B)) = \text{supp}(c(c(\pi, B), A))$ . This is because



if the support of the posteriors under  $A$  and  $B$  intersect, then  $A$  and  $B$  both must be intersecting with the same equivalence class of  $\geq_\pi$ .

It can be shown that lexicographic updating rules satisfy weak path independence. Suppose  $\{\mathbf{q}_\pi\}$  is the associated family of LPS's. Since  $\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset$ , it follows that  $\text{supp}(\mathbf{q}^\pi(\cdot|A)) \cap \text{supp}(\mathbf{q}^\pi(\cdot|B)) \neq \emptyset$ . Hence, we have  $\min\{k : q_k^\pi(A) > 0\} = \min\{k : q_k^\pi(B) > 0\} =: k^*$ . Moreover,  $A \cap B \cap \text{supp}(q_{k^*}^\pi) \neq \emptyset$ . This means that  $A \cap B$  is a positive probability event under  $q_{k^*}^\pi$ . Then, by Bayes' rule applied to the probability measure  $q_{k^*}^\pi$  and events  $A$  and  $B$ , we get that  $c(c(\pi, A), B) = q_{k^*}^\pi(\cdot|A \cap B) = c(c(\pi, B), A)$ . Hence, the updating rules satisfies weak path independence. We now prove the main result.

**Proposition 3.** *Let  $\langle c, \Pi \rangle$  be an updating rule which is Bayesian, AGM-consistent and satisfies weak path independence. Then  $\langle c, \Pi \rangle$  is lexicographic.*

*Proof.* We wish to define a family  $\{\mathbf{q}^\pi\}_{\pi \in \Delta(\Omega)}$  of LPS's. We do this as follows. Let  $\pi \in \Delta(\Omega)$ . Since  $\langle c, \Pi \rangle$  is AGM-consistent, there exists a complete transitive relation  $\geq_\pi$  on  $\Omega$  which determines  $\Pi(\pi, \cdot)$ . Now define  $\mathbf{q}^\pi$  in the following way :

$$\text{Set } q_k^\pi = c(\pi, M_{\geq_\pi}^k(\Omega)) \text{ for all } k. \quad (3)$$

Here, recall that  $M_{\geq_\pi}^k(\Omega)$  is the set of all  $k$ -th highest states (equivalence class) according to  $\geq_\pi$ . Define  $\mathbf{q}^\pi = \langle q_k^\pi \rangle_k$ . Note that the support of  $q_k^\pi$  is  $M_{\geq_\pi}^k(\Omega)$ . This is true since  $c$  is AGM-consistent. The support of  $c(\pi, A)$  is precisely  $M_{\geq_\pi}(A)$ , which is the set of all maximal elements of  $A$  according to  $\geq_\pi$ . But now note that  $M_{\geq_\pi}(M_{\geq_\pi}^k(\Omega)) = M_{\geq_\pi}^k(\Omega)$ , since all states in the equivalence class  $M_{\geq_\pi}^k(\Omega)$  are equally ranked according to  $\geq_\pi$ . Hence, the support of  $q_k^\pi$  is  $M_{\geq_\pi}^k(\Omega)$  from the definition of  $q_k^\pi$  in 3.

Now let  $A \subseteq \Omega$  be non-empty. Let  $k(\pi, A) = \min\{k : q_k^\pi(A) > 0\}$ . So now we get that  $M_{\geq_\pi}(A) = M_{\geq_\pi}^{k(\pi, A)}(\Omega) \cap A = \text{supp}(q_{k(\pi, A)}^\pi) \cap A = \text{supp}(\mathbf{q}^\pi(\cdot|A))$ . Hence, we get  $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = M_{\geq_\pi}(A)\} = \{\pi' : \text{supp}(\pi') = \text{supp}(\mathbf{q}^\pi(\cdot|A))\}$ . Also :

$$\begin{aligned} \mathbf{q}^\pi(\cdot|A) &= q_{k(\pi, A)}^\pi(\cdot|A) \\ &= c(q_{k(\pi, A)}^\pi, A) \\ &= c(c(\pi, M_{\geq_\pi}^{k(\pi, A)}(\Omega)), A) \end{aligned} \quad (4)$$

$$= c(c(\pi, A), M_{\geq_\pi}^{k(\pi, A)}(\Omega)) \quad (5)$$

$$= c(\pi, A) \quad (6)$$

where 4 and 6 follow from the fact that  $\langle c, \Pi \rangle$  is Bayesian and 5 follows from weak path

independence. Hence, we have shown that  $\langle c, \Pi \rangle$  is lexicographic.  $\square$

Combining the results above we obtain the following.

**Corollary 4.** *An updating rule is Bayesian, AGM-consistent and satisfies weak path independence if and only if it is lexicographic.*

We now show that the properties "AGM-consistency", "Bayesian" and "Weak path independence" are independent from each other i.e. there exist updating rules which satisfy two of the properties but not the third. For convenience, let us call them AGM, BA and WPI respectively. We demonstrate the independence of the three properties via the following examples.

**Example 3.1.** (*AGM, BA but not WPI*) Let  $\Omega = \{x, y, z, w\}$  and let  $\langle c, \Pi \rangle$  be any Bayesian updating rule such that  $\Pi(\pi, A) = \{\pi' | \text{supp}(\pi') = A\}$  whenever  $\pi(A) = 0$ . One can verify that this is AGM-consistent. For each  $\pi$ , define  $\geq_\pi$  as a complete and transitive relation with two equivalence classes, namely  $\text{supp}(\pi)$  and its complement, where the former is ranked higher than the latter. Now suppose,  $c(\delta_x, \{y, z\}) = 1/2(y) + 1/2(z) = (0, 1/2, 1/2, 0)$  and  $c(\delta_x, \{y, z, w\}) = 1/2(y) + (1/4)(z) + 1/4(w)$ . Now let  $\pi = \delta_x, A = \{y, z\}, B = \{y, z, w\}$ . Then,  $c(c(\pi, A), B) = (0, 1/2, 1/2, 0)$  but  $c(c(\pi, B), A) = (0, 2/3, 1/3, 0)$ , violating weak path independence. From Proposition 3, this also serves as an example of a Bayesian and AGM-consistent updating rule that is not lexicographic.

**Example 3.2.** (*AGM, WPI but not BA*) Let  $\Omega = \{x, y, z\}$  and let  $\langle c, \Pi \rangle$  be an updating rule which always picks the posterior which has uniform distribution on the event observed  $A$ . Since we are always updating to a uniform distribution, weak path independence is trivially satisfied. AGM consistency is also satisfied as in the previous example. Note however, that such an updating rule cannot be Bayesian since for  $\pi = (1/2, 1/3, 1/6)$ , and  $A = \{x, y\}$ , we get  $c(\pi, A) = (1/2, 1/2, 0)$  which is not the Bayesian posterior.

**Example 3.3.** (*WPI, BA but not AGM*) Let  $\langle c, \Pi \rangle$  be any Bayesian updating rule such that  $\text{supp}\Pi(\pi, A)$  is a singleton whenever  $\pi(A) > 0$ , meaning that  $c(\pi, A)$  is a delta measure. We can show that such rules always satisfy weak path independence. Let  $\pi \in \Delta(\Omega)$  and  $A, B \subseteq \Omega$  such that  $\text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) \neq \emptyset$ .

- Case 1: W.l.o.g suppose  $\pi(A) > 0$ . Then  $\emptyset \neq \text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) = \text{supp}(\pi) \cap A \cap \text{supp}\Pi(\pi, B) \subseteq \text{supp}(\pi) \cap A \cap B \subseteq A \cap B$ . This implies that  $\pi(A \cap B) > 0$ . Since  $\langle c, \Pi \rangle$  is Bayesian, we have  $c(c(\pi, A), B) = c(c(\pi, B), A)$  because restricted to positive probability events, Bayes' rule guarantees path independence.

- Case 2 :  $\pi(A) = 0$  and  $\pi(B) = 0$ . By assumption,  $\emptyset \neq \text{supp}\Pi(\pi, A) \cap \text{supp}\Pi(\pi, B) = \text{supp}\Pi(\pi, A) = \text{supp}\Pi(\pi, B) = \{\omega\}$ , for some  $\omega \in A \cap B$ . Hence,  $c(\pi, A) = c(\pi, B) = \delta_\omega$ . Since  $\langle c, \Pi \rangle$  is Bayesian,  $c(c(\pi, A), B) = c(c(\pi, B), A) = \delta_\omega$ .

Now, let  $\Omega = \{x, y, z, w\}$  and consider any such rule  $\langle c, \Pi \rangle$  which satisfies  $c(\delta_x, \{y, z, w\}) = \delta_y$  and  $c(\delta_x, \{y, z\}) = \delta_z$ . Such rules exist because the only restriction here for the case  $\pi(A) = 0$  is that the support of  $c(\pi, A)$  is a singleton. Suppose for contradiction that  $\langle c, \Pi \rangle$  is AGM-consistent and let  $\{\geq_\pi\}_\pi$  be the associated family of complete and transitive orderings. Since,  $c(\delta_x, \{y, z, w\}) = \delta_y$ , we must have that  $y \geq_{\delta_x} z$  and  $\neg z \geq_{\delta_x} y$ . However, on the other hand, since  $c(\delta_x, \{y, z\}) = \delta_z$ , we must have  $z \geq_{\delta_x} y$ , which is a contradiction.

## 4 Minimum Distance Updating Rules

We now study minimum distance updating rules. Such updating rules select posteriors in  $\Pi(\pi, A)$  which are "closest" to the prior.

**Definition 4.1.** An updating rule  $\langle c, \Pi \rangle$  is said to be *minimum distance* if there exists a metric  $d$  on  $\Delta(\Omega)$  such that  $\{c(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ .

Minimum distance updating rules capture the idea of conservatism in belief updating, in the sense that new information keeps updated beliefs as close to the prior belief as is possible. Not all Bayesian updating rules have a minimum distance representation. However, certain restrictions on the AGM procedure allow Bayesian updating rules to be minimum distance.

We explore the connection between lexicographic and minimum distance updating rules. We show that a sub-class of lexicographic updating rules, which we call "support dependent" updating rules necessarily have a minimum distance representation. Additionally, we demonstrate an example of a lexicographic updating rule outside this class which has no minimum distance representation.

**Definition 4.2.** A lexicographic updating rule  $\langle c, \Pi \rangle$  is *support dependent* if the family of lexicographic updating rules  $\{\mathbf{p}^\pi\}_\pi$  generating it satisfies :

$$\text{For all } \pi, \pi' \in \Delta(\Omega), \text{supp}(\pi) = \text{supp}(\pi') \text{ implies } \mathbf{p}^\pi \setminus \{\pi\} = \mathbf{p}^{\pi'} \setminus \{\pi'\}$$

The above condition is interpreted as follows : If two decision makers have priors with the same support, then they use the same secondary hypotheses to update when they receive surprising information. The next result shows that support dependent lexicographic updating rules have a minimum distance representation.

**Proposition 5.** *Every support dependent lexicographic updating rule has a minimum distance representation.*

*Proof.* The proof can be found in the appendix and involves two main steps. We provide here an outline of the proof. In the first step, we construct a metric  $d$  that gaurantees that  $c(\pi, A)$  is the unique solution to  $\min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$  whenever  $\pi(A) > 0$ . Hence, essentially  $d$  gaurantees a minimum distance representation of Bayesian updating on positive probability events. In the second step, we use  $d$  to construct a metric  $d_L$  which provides a minimum distance representation of  $\langle c, \Pi \rangle$ . For any  $\pi, \pi'$ , the quantity  $d_L(\pi, \pi')$  is defined as a translation of the sum of the distances  $d(p_l^\pi, \pi')$  and  $d(\pi, p_k^{\pi'})$  where  $l$  is the smallest integer so that  $\text{supp}(p_l^\pi) \cap \text{supp}(\pi') \neq \emptyset$  and  $k$  is the smallest integer so that  $\text{supp}(p_k^{\pi'}) \cap \text{supp}(\pi) \neq \emptyset$ . This covers the case  $\pi(A) = 0$ .  $\square$

The following is an example of a lexicographic updating rule that is not support dependent and admits no minimum distance representation.

**Example 4.1.** Let  $\Omega = \{x, y, z, w\}$ . A typical probability vector in  $\Delta(\Omega)$  will be denoted as  $\pi = (\pi_x, \pi_y, \pi_z, \pi_w)$ . For each  $A \subseteq \Omega$ , we denote  $\pi_A^u \in \Delta(\Omega)$  to be the uniform distribution supported on  $A$ . Define the following family  $\{\mathbf{p}^\pi\}_\pi$  of LPS's

$$\mathbf{p}^\pi = \begin{cases} \langle (p, 1-p, 0, 0), (0, 0, p, 1-p) \rangle & \text{if } \pi = (p, 1-p, 0, 0) \\ \langle (0, 0, p, 1-p), (1-p, p, 0, 0) \rangle & \text{if } \pi = (0, 0, p, 1-p) \\ \langle \pi, \pi_{\text{supp}(\pi)^c}^u \rangle & \text{o.w.} \end{cases}$$

Note that  $\mathbf{p}^\pi$  is not support dependent. Suppose for contradiction that the lexicographic updating rule generated by the family  $\{\mathbf{p}^\pi\}_\pi$  has a minimum distance representation. Let  $d$  be the associated metric.

Let  $A = \{x, y\}$  and  $B = \{z, w\}$ . Firstly, note that  $c((1/4, 3/4, 0, 0), B) = (0, 0, 1/4, 3/4)$  ;  $c((0, 0, 1/4, 3/4), A) = (3/4, 1/4, 0, 0)$  ;  $c((3/4, 1/4, 0, 0), B) = (0, 0, 3/4, 1/4)$  ;  $c((0, 0, 3/4, 1/4), A) = (1/4, 3/4, 0, 0)$ . Hence, starting with prior  $(1/4, 3/4, 0, 0)$  and alternating events  $B$  and  $A$  we obtain a cycle back to  $(1/4, 3/4, 0, 0)$ . Since  $c(\pi, A)$  is the unique minimiser of distance, we get the following chain of strict inequalities

$$\begin{aligned} d(((1/4, 3/4, 0, 0), (0, 0, 1/4, 3/4)) &< d(((1/4, 3/4, 0, 0), (0, 0, 3/4, 1/4)) \\ &< d(((3/4, 1/4, 0, 0), (0, 0, 3/4, 1/4)) \\ &< d(((3/4, 1/4, 0, 0), (0, 0, 1/4, 3/4)) \\ &< d(((1/4, 3/4, 0, 0), (0, 0, 1/4, 3/4)). \end{aligned}$$

This results in a contradiction. Hence, the updating rule corresponding to  $\{\mathbf{p}^\pi\}_\pi$  has no minimum distance representation.

It is interesting to compare the present setting involving minimum distance updating rules with Perea [2009]. In that paper, updating of  $\pi$  based on  $E$  involves choosing the minimiser of the euclidean distance  $\|\phi(\pi) - \phi(\pi')\|$  over  $\pi' \in \Delta(E)$ , where  $\phi$  is an affine map and  $\Delta(E)$  is the set of all probability measures that have support in  $E$ . The approach of the paper is related to the present paper though the notions of minimum distance and the nature of updating is very different. For example, Perea [2009] is motivated by the idea of imaging and non-Bayesian updating. Also, the updating rules need not be AGM consistent. Lastly, Rubinstein and Zhou [1999] also discuss choice problems where alternatives are chosen according to the minimum distance rule. One may view the present work as being complementary to these papers.

## 5 Updating Preferences

It is possible to undertake a similar approach to the problem of updating preferences (see for example Gilboa and Schmeidler [1993] and Hanany and Klibanoff [2007]). Let  $\Omega$  be the state space as before and let  $X$  denote a finite set of *outcomes*. An *act* is a map  $f : \Omega \rightarrow \Delta(X)$  and let  $\mathcal{F}$  denote the set of all acts. A preference relation  $\succsim$ , is any complete and transitive relation on  $\mathcal{F}$ . We let  $\mathcal{P}$  denote any collection of preferences.

An updating rule in this context is defined as follows. The agent starts with a prior preference  $\succsim$  and then upon observing  $A$ , updates his preference to  $\succsim'$ . Formally, an updating rule is defined analogously as a tuple  $\langle c, \Pi \rangle$ , where  $c : \mathcal{P} \times 2^\Omega \setminus \{\emptyset\} \rightarrow \mathcal{P}$  and  $\Pi : \mathcal{P} \times 2^\Omega \setminus \{\emptyset\} \rightarrow 2^\mathcal{P}$ . As before, the idea is that when the preference is  $\succsim$  and the event observed is  $A$ , then the agent first narrows down a menu of possible preferences  $\Pi(\succsim, A)$ , which corresponds to the qualitative part of updating, possibly involving identifying a set of core beliefs (belief sets below). Next, the choice  $c(\succsim, A)$  corresponds to the quantitative part of updating, which involves the precise pairwise comparisons of various acts, depending on the lotteries over outcomes that would yield at each of the states.

In this context, one may define belief sets corresponding to preferences. Firstly, we define conditional preferences in the sense of Savage. For any  $A$ , we say  $f \succsim_A g$  if  $f_A h \succsim g_A h$  for all  $h$  (see Savage [1954]). Here, for two acts  $f$  and  $h$  and an event  $A$ , the act  $f_A h$  is defined as the act that agrees with  $f$  on  $A$  and agrees with  $h$  on  $A^c$ . If the preference relation

$\succsim$  satisfies the Independence axiom<sup>4</sup>, it is straightforward to show that the relation  $\succsim_E$  is complete, transitive and also satisfies the Independence axiom. For each preference  $\succsim$ , we say an event  $E$  is  $\succsim$ -savage-null if  $f_E h \sim g_E h$  for all  $h$ .

**Definition 5.1.** A *belief set* corresponding to a preference relation  $\succsim$  is defined as the set of events

$$\mathcal{K}(\succsim) = \{E : E^c \text{ is } \succsim\text{-savage-null}\}. \quad (7)$$

Hence, a  $\succsim$ -savage-null event corresponds to a set of states that are essentially disregarded in the agent's preference. It is straightforward to show that  $\mathcal{K}(\succsim)$  is always logically closed. This follows from the fact that the union of two  $\succsim$ -savage-null events is also  $\succsim$ -savage-null and the fact that subsets of  $\succsim$ -savage-null events are also  $\succsim$ -savage-null. Note that when  $E$  is  $\succsim$ -savage-null, then the relation  $\succsim_E$  is trivial i.e.  $f \sim_E g$  for all  $f$  and  $g$ . Also, when  $A$  is such that  $\cap_{E \in \mathcal{K}(\succsim)} E \cap A \neq \emptyset$ , we have that  $\cap_{E \in \mathcal{K}(\succsim_A)} E = \cap_{E \in \mathcal{K}(\succsim)} E \cap A$ . It shall be convenient to define the set  $B(\succsim) = \cap_{E \in \mathcal{K}(\succsim)} E$  for  $\succsim \in \mathcal{P}$ , corresponds to the "core" of the belief set  $\mathcal{K}(\succsim)$ .

We will require that updating rules satisfy the condition that  $B(\succsim') = B(\succsim'') \subseteq A$  for all  $\succsim', \succsim'' \in \Pi(\succsim, A)$  i.e. the belief set corresponding to the updated preferences contains the arrived information  $A$ . The common support, as before, will be denoted as  $\text{supp}\Pi(\succsim, A)$ . We say an updating rule is *savage-consistent* if  $c(\succsim, A) = \succsim_A$  whenever  $B(\succsim) \cap A \neq \emptyset$  i.e. whenever  $A$  is consistent with the belief set  $\mathcal{K}(\succsim)$ . An updating rule is said to be *AGM-consistent* if there exists a family of revision operators  $\{*_\succsim\}_{\succsim \in \mathcal{P}}$  such that

$$\Pi(\succsim, A) = \{\succsim' \in \mathcal{P} : \mathcal{K}(\succsim') = \mathcal{K}(\succsim) *_\succsim A\}. \quad (8)$$

The notions of strong and weak path independence arise analogously. An updating rule satisfies *strong path independence* if for all  $A \cap B \neq \emptyset$ , we have  $c(c(\succsim, A), B) = c(c(\succsim, B), A)$ . It satisfies *weak path independence* if  $c(c(\succsim, A), B) = c(c(\succsim, B), A)$  is required to hold whenever  $\text{supp}\Pi(\succsim, A) \cap \text{supp}\Pi(\succsim, B) \neq \emptyset$ .

In line with Proposition 2 and Proposition 3, we obtain here a characterisation of *lexicographic* updating rules, which are defined as follows. A *lexicographic preference profile* is a finite sequence of preference relations  $\mathbf{P} = \langle P_1, \dots, P_K \rangle$ <sup>5</sup> such that  $\cup_k B(P_k) = \Omega$  and  $B(P_k) \cap B(P_l) = \emptyset$  for all  $k \neq l$ . An updating rule  $\langle c, \Pi \rangle$  is said to be *lexicographic* if

<sup>4</sup>The Independence axiom states that for all  $f, g, h$  and for all  $\lambda \in (0, 1)$ , it holds that  $f \succsim g$  if and only if  $\lambda f + (1 - \lambda)h \succsim \lambda g + (1 - \lambda)h$ .

<sup>5</sup>We adopt the notation  $P_k$  to differentiate it with the prior preference  $\succsim$ .

there exists a family a lexicographic preference profiles  $\{\mathbf{P}^{\succsim}\}_{\succsim \in \mathcal{P}}$  such that  $\succsim = P_1^{\succsim}$  and

$$c(\succsim, A) = P_{k,E}^{\succsim}, \quad (9)$$

where  $k = \min\{l : B(P_l^{\succsim}) \cap A \neq \emptyset\}$  and  $P_{k,E}^{\succsim}$  is the savage conditional preference based on the event  $E$ . The following holds in this context.

**Proposition 6.** *Let  $\mathcal{P}$  be a class of preference relations which satisfy the independence axiom. Suppose that for each  $\succsim \in \mathcal{P}$ , it holds that  $\succsim_A \in \mathcal{P}$  whenever  $B(\succsim) \cap A \neq \emptyset$ . Further, suppose that  $B(\succsim) \neq \emptyset$  for each  $\succsim \in \mathcal{P}$  and that for each non-empty  $A$ , there exists  $\succsim \in \mathcal{P}$  such that  $B(\succsim) = A$ . Then, the following hold.*

1. *If  $|\Omega| \geq 3$ , there does not exist any updating rule that is both Savage-consistent and satisfies strong path independence.*
2. *An updating rule is Savage-consistent, AGM-consistent and satisfies weak path independence if and only if it is lexicographic.*

Minimum distance rules may be defined in the same manner as in Section 4. This would involve defining a metric over preference relations,  $d(\succsim, \succsim')$ . The updating rule would have  $c(\pi, A)$  be the unique solution to the problem  $\min_{\succsim' \in \Pi(\succsim, A)} d(\succsim, \succsim')$ . It could potentially be useful to study such updating rules and its properties as an extension of the present work.

## 6 Conclusion

In this paper, we studied the problem of updating over zero probability events. Our analysis is centered around the relationship between the problem of theory change in the belief revision literature in propositional logic and updating probabilistic beliefs about events based on evidence. In particular, we studied lexicographic and minimum-distance updating rules and their relationship with AGM consistent updating rules. As noted earlier in the introduction, there have been alternative approaches to this problem and one may also consider extensions of our analysis. One extension, as indicated in the introduced, would be to study implications of our approach to updating ambiguous beliefs. Another line of extension would be to investigate how our approach may apply to studying equilibrium refinements and issues of unawareness in games.



## References

- C. E. Alchourron, P. Gardenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530., 1985.
- Pierpaolo Battigalli and Marciano Siniscalchi. Strong belief and forward induction reasoning. *Journal of Economic Theory*, 106(2):356–391, 2002.
- L. Blume, A. Brandenburger, and E. Dekel. Lexicographic probabilities and choice under uncertainty. *Econometrica*, 59:61–79, 1991a.
- Giacomo. Bonanno. Rational choice and agm belief revision. *Artificial Intelligence*, 173: 1194–1203, 2009.
- H. Costa and A. Pedersen. Belief revision. *The Continuum Companion to Philosophical Logic*, pages 450–502, 2011.
- Jean H Gallier. *Logic for computer science: foundations of automatic theorem proving*. Courier Dover Publications, 2015.
- I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *J. Math. Econ*, 18:141–153, 1989.
- I. Gilboa and D. Schmeidler. Updating ambiguous beliefs. *Journal of Economic Theory*, 59: 33–49, 1993.
- Adam Grove. Two modellings for theory change. *Journal of philosophical logic*, 17(2): 157–170, 1988.
- J. Halpern. Lexicographic probability, conditional probability, and nonstandard probability. *Games and Economic Behavior*, 68:1:155–179, 2010.
- P. J. Hammond. Elementary non-archimedean representations of probability for decision theory and games. *P. Humphreys (Ed.), Patrick Suppes: Scientific Philosopher*, 1:25–49, 1994.
- Eran Hanany and Peter Klibanoff. Updating preferences with multiple priors. *Theoretical Economics*, 2(3):261–298, 2007.
- Franz Huber. Belief revision i: the agm theory. *Philosophy Compass*, 8(7):604–612, 2013.
- Edi Karni and Marie-Louise Vierø. ” reverse bayesianism”: A choice-based theory of growing awareness. *American Economic Review*, 103(7):2790–2810, 2013.

- D. M. Kreps and R. B. Wilson. Sequential equilibria. *Econometrica*, 50:863–894, 1982b.
- Roger B Myerson. Axiomatic foundations of bayesian decision theory. *Center for Mathematical Studies in Economics and Management Science, Northwestern University Working Paper*, 671, 1986.
- Pietro Ortoleva. Modeling the change of paradigm: Non-bayesian reactions to unexpected news. *American Economic Review*, 102(6):2410–36, 2012.
- Andrés Perea. A model of minimal probabilistic belief revision. *Theory and Decision*, 67(2):163–222, 2009.
- A. Rubinstein and L. Zhou. Choice problems with a "reference" point. *Mathematics of Social Sciences*, 37:205–210, 1999.
- L. J Savage. The foundations of statistics. 1954.
- D. Schmeidler. Subjective probability and expected utility without additivity. *Econometrica*, 57:571–587, 1989.
- R. Selten. Re-examination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*, 4:25–55, 1975.
- Y. Shoham and .K Leyton-Brown. Multiagent systems: Algorithmic, game-theoretic and logical foundations. *Cambridge University Press*, 2009.

## 7 Appendix

### 7.1 Proof of Proposition 5

*Proof. Step 1 :*

We define the metric  $d$  on  $\Delta(\Omega)$ . First, for any pair of states  $\omega, \omega' \in \Omega$ , we define the function  $S_{\omega\omega'} : \Delta(\Omega) \times \Delta(\Omega) \rightarrow \mathbb{R}_+$  as follows :

- If  $\pi \neq \pi'$  and  $\min\{\pi(\{\omega, \omega'\}), \pi'(\{\omega, \omega'\})\} > 0$  :

$$S_{\omega\omega'}(\pi, \pi') = \left| \frac{\pi(\omega)}{\pi(\{\omega, \omega'\})} - \frac{\pi'(\omega)}{\pi'(\{\omega, \omega'\})} \right|.$$

- If  $\pi \neq \pi'$  and  $\min\{\pi(\{\omega, \omega'\}), \pi'(\{\omega, \omega'\})\} = 0$  :

$$S_{\omega\omega'}(\pi, \pi') = c \text{ where } c \geq 1.$$

- If  $\pi = \pi'$  then :

$$S_{\omega\omega'}(\pi, \pi') = 0.$$

Notice that for any pair  $\omega, \omega' \in \Omega$ ,  $S_{\omega\omega'}(\pi, \pi') = S_{\omega'\omega}(\pi, \pi')$ . Also, it can easily be shown that  $S_{\omega\omega'}$  defines a pseudometric on  $\Delta(\Omega)$  and that for any  $\pi \neq \pi'$ , there exist states  $\omega, \omega' \in \Omega$  such that  $S_{\omega\omega'}(\pi, \pi') > 0$ . Combining these two facts we get that the following is a metric on  $\Delta(\Omega)$ .

$$d(\pi, \pi') = \sum_{\omega \neq \omega'} S_{\omega\omega'}(\pi, \pi').$$

The above metric is indeed exactly the one we need. So we shall now show that for any  $A \subseteq \Omega$  such that  $\pi(A) > 0$ , we have  $c(\pi, A) = \{\pi_B(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ . Here,  $\pi_B(\pi, A)$  denotes the Bayesian update of  $\pi$  based on  $A$ . Now, note if  $\text{supp}(\pi) \cap A = \{\omega\}$ , then  $\Pi(\pi, A) = \{\delta_\omega\}$ , so clearly  $\{\pi_B(\pi, A)\} = \{\delta_\omega\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ . So we only consider the case when  $|\text{supp}(\pi) \cap A| \geq 2$ .

- *Case 1* :  $\pi \in \Pi(\pi, A)$  :

Notice in this case,  $\text{supp}(\pi) = \text{supp}(\pi) \cap A$ , hence  $\pi(A) = 1$ . So we have  $\pi = \pi_B(\pi, A)$ . Now clearly,  $d(\pi, \pi_B(\pi, A)) = 0$ . Now consider  $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$ . Since  $\pi \neq \pi_B(\pi, A) = \pi$ , there exists a pair of states  $\omega_1, \omega_2 \in \Omega$  such that  $S_{\omega_1\omega_2}(\pi, \pi') > 0$ . Hence  $d(\pi, \pi') > 0$  and we get  $\{\pi_B(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ .

- *Case 2* :  $\pi \notin \Pi(\pi, A)$  :

Now partition pairs of states in the following way :

- $S^o = \{(\omega, \omega') : \omega \in \text{supp}(\pi) \cap A \text{ but } \omega' \notin \text{supp}(\pi) \cap A\}$ .
- $S^- = \{(\omega, \omega') : \{\omega, \omega'\} \subseteq [\text{supp}(\pi) \cap A]^c \text{ and } \omega \neq \omega'\}$ .
- $S^+ = \{(\omega, \omega') : \{\omega, \omega'\} \subseteq \text{supp}(\pi) \cap A \text{ and } \omega \neq \omega'\}$ .

1. Consider  $(\omega, \omega') \in S^o$ . Since  $\omega \in \text{supp}(\pi) \cap A$ ,  $\pi(\{\omega, \omega'\}) > 0$  and  $\pi'(\{\omega, \omega'\}) > 0$  for all  $\pi' \in \Pi(\pi, A)$ . Hence,  $\min\{\pi(\{\omega, \omega'\}), \pi'(\{\omega, \omega'\})\} > 0$  and  $\pi \neq \pi'$  for all  $\pi' \in \Pi(\pi, A)$  (since  $\pi \notin \Pi(\pi, A)$ ). Hence :

$$S_{\omega\omega'}(\pi, \pi') = \left| \frac{\pi(\omega)}{\pi(\{\omega, \omega'\})} - \frac{\pi'(\omega)}{\pi'(\{\omega, \omega'\})} \right| = \left| \frac{\pi(\omega)}{\pi(\{\omega, \omega'\})} - 1 \right| \text{ for all } \pi' \in \Pi(\pi, A).$$

So  $S_{\omega\omega'}(\pi, \pi') = S_{\omega\omega'}(\pi, \pi'')$  for all  $\pi', \pi'' \in \Pi(\pi, A)$ . As a result we get

$$\sum_{(\omega, \omega') \in S^0} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = \sum_{(\omega, \omega') \in S^0} S_{\omega, \omega'}(\pi, \pi').$$

for all  $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$ .

2. Consider  $(\omega, \omega') \in S^-$ . Then clearly  $\pi'(\{\omega, \omega'\}) = 0$  and  $\pi \neq \pi'$  for all  $\pi' \in \Pi(\pi, A)$ . Hence,  $S_{\omega, \omega'}(\pi, \pi') = c$  for all  $\pi' \in \Pi(\pi, A)$ . As a result we get

$$\sum_{(\omega, \omega') \in S^-} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = \sum_{(\omega, \omega') \in S^-} S_{\omega, \omega'}(\pi, \pi') \text{ for all } \pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}.$$

3. Consider  $(\omega, \omega') \in S^+$ . Then clearly  $\pi(\{\omega, \omega'\}) > 0$  and  $\pi'(\{\omega, \omega'\}) > 0$  for all  $\pi' \in \Pi(\pi, A)$ . Hence,  $\min\{\pi(\{\omega, \omega'\}), \pi'(\{\omega, \omega'\})\} > 0$  and  $\pi \neq \pi'$  for all  $\pi' \in \Pi(\pi, A)$ . Notice that  $S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = 0$  since  $\frac{\pi_B(\pi, A)(\omega)}{\pi_B(\pi, A)(\{\omega, \omega'\})} = \frac{\frac{\pi(\omega)}{\pi(A)}}{\frac{\pi(\{\omega, \omega'\})}{\pi(A)}} = \frac{\pi(\omega)}{\pi(\{\omega, \omega'\})}$ . So we get  $S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) = 0$  for all  $(\omega, \omega') \in S^+$ .

Now consider  $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$ . Now there exist states  $\omega, \omega' \in E * (\pi) \cap A$  such that  $S_{\omega, \omega'}(\pi, \pi') > 0$ .

As a result we get  $\sum_{(\omega, \omega') \in S^+} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) < \sum_{(\omega, \omega') \in S^+} S_{\omega, \omega'}(\pi, \pi')$  for all  $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$ .

Combining the results of 1 2 and 3 above, we get  $d(\pi, \pi_B(\pi, A)) = \sum_{\omega \neq \omega'} S_{\omega, \omega'}(\pi, \pi_B(\pi, A)) <$

$\sum_{\omega \neq \omega'} S_{\omega, \omega'}(\pi, \pi') = d(\pi, \pi')$  for all  $\pi' \in \Pi(\pi, A) \setminus \{\pi_B(\pi, A)\}$ . Hence, we obtain our desired result that  $\{\pi_B(\pi, A)\} = \arg \min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ .

## Step 2 :

Let  $< c, \Pi >$  be a support dependent lexicographic updating rule and let  $\{\mathbf{p}^\pi\}_\pi$  be the family generating it. Consider the metric  $d$  constructed in Step 1. Now we define a function  $\hat{d} : \Delta(\Omega) \times \Delta(\Omega) \rightarrow \mathbb{R}_+$  as follows :

1. For  $\pi = \pi'$ , set  $\hat{d}(\pi, \pi') = 0$ .
2. For  $\pi \neq \pi'$ , let  $\mathbf{p}^\pi = \{p_k^\pi\}_k$  and  $\mathbf{p}^{\pi'} = \{p_l^{\pi'}\}_l$  define the natural number  $K(\pi, \pi')$  as :

$$K(\pi, \pi') := \min\{l : \text{supp}(p_l^\pi) \cap \text{supp}(\pi') \neq \emptyset\}$$

And now define :

$$\hat{d}(\pi, \pi') = d(p_{K(\pi, \pi')}^\pi, \pi').$$

The interpretation of  $\hat{d}$  is as follows. Given the LPS  $\mathbf{p}^\pi = (p_k^\pi)_k$  corresponding to  $\pi$ , we consider at the smallest  $l$  such that  $\text{supp}(p_l^\pi) \cap \text{supp}(\pi') \neq \emptyset$  i.e. such that  $\pi'$  is consistent with  $\mathbf{p}^\pi$ . Hence, in a sense, the function  $\hat{d}(\pi, \pi')$  measures the "divergence" between  $\pi$  and  $\pi'$  in terms of the LPS  $\mathbf{p}^\pi$  and in particular, the secondary hypotheses in  $\mathbf{p}^\pi$  are applied for the case when  $\text{supp}(\pi) \cap \text{supp}(\pi') = \emptyset$ . Note further that in the case  $\text{supp}(\pi) \cap \text{supp}(\pi') = \emptyset$ , we would have that  $d(\pi, \pi') = c$  from the construction of  $d$  in Step 1 above. Also, in the case when  $\text{supp}(\pi) \cap \text{supp}(\pi') \neq \emptyset$ , it follows that  $d(\pi, \pi') = \hat{d}(\pi, \pi')$ . Hence, the function  $\hat{d}(\pi, \pi')$  distinguishes between  $\pi$  and  $\pi'$  more "finely" compared to  $d$ .

The function  $\hat{d}(\pi, \pi')$  will now be used to define our the desired metric  $d_L$ . Note  $\hat{d}(\pi, \pi')$  itself may not be symmetric or satisfy the triangular inequality. We proceed as follows. First, we define the function  $\hat{S}(\pi, \pi') = \hat{d}(\pi, \pi') + \hat{d}(\pi', \pi)$ . Note now that  $\hat{S}$  is symmetric and  $0 < \sup_{\pi, \pi'} \hat{S}(\pi, \pi') < +\infty$  but may not satisfy triangular inequality. Now, let  $M > 0$  be a real number such that  $\sup_{\pi, \pi'} \hat{S}(\pi, \pi') < M$ . Finally, define the desired metric  $d_L$  as follows :

1. For  $\pi = \pi'$ , set  $d_L(\pi, \pi') = 0$ .
2. For  $\pi \neq \pi'$ , set  $d_L(\pi, \pi') = \hat{S}(\pi, \pi') + M$ .

It can be checked that  $d_L$  satisfies the axioms of a metric and is essentially a translation of  $\hat{S}$  so as to satisfy the triangular inequality.

We now show that the above metric  $d_L$  generates  $\langle c, \Pi \rangle$ . Note that since  $M$  is a constant, the solutions to the minimisation problem  $\min_{\pi' \in \Pi(\pi, A)} d_L(\pi, \pi')$  and the problem  $\min_{\pi' \in \Pi(\pi, A)} \hat{S}(\pi, \pi')$  are exactly the same. Hence, we will show that the unique solution is equal to  $c(\pi, A)$ .

Let  $\pi \in \Delta(\Omega)$  and  $A \in 2^\Omega \setminus \{\emptyset\}$  and define  $k(\pi, A) = \min\{k : p_k^\pi(A) > 0\}$ . Also, recall that  $\Pi(\pi, A) = \{\pi' : \text{supp}(\pi') = \text{supp}(p_k^\pi) \cap A\}$ . There are two cases.

1. Suppose first that  $\pi(A) > 0$ . Then,  $d_L(\pi, \pi') = 2d(\pi, \pi') + M$  for all  $\pi' \in \Pi(\pi, A)$ . From Step 1, we indeed have that  $c(\pi, A)$ , which is the Bayesian update of  $\pi$  based on  $A$ , is the unique solution to the minimisation problem  $\min_{\pi' \in \Pi(\pi, A)} d(\pi, \pi')$ . This in turn implies that  $c(\pi, A)$  is also the unique solution to  $\min_{\pi' \in \Pi(\pi, A)} d_L(\pi, \pi')$ .
2. Now suppose that  $\pi(A) = 0$ . Note that this means that  $k(\pi, A) \geq 2$ . Recall all probability measures in  $\Pi(\pi, A)$  have the same support, which in this case is disjoint from the support of  $\pi$ . Hence, from support dependence of the updating rule  $\langle c, \Pi \rangle$ , we

have that  $\hat{d}(\pi', \pi) = \hat{d}(\pi'', \pi)$  for all  $\pi', \pi'' \in \Pi(\pi, A)$ . This follows directly from the definition of  $\hat{d}$ .

Also from the definition of  $\hat{d}$  in terms of  $d$  and from Step 1, we have that  $\hat{d}(\pi, \pi') = d(p_{k(\pi, A)}^\pi, \pi') > d(p_{k(\pi, A)}^\pi, p_{k(\pi, A)}^\pi(\cdot|A)) = \hat{d}(\pi, c(\pi, A))$  for all  $\pi' \neq c(\pi, A)$  in  $\Pi(\pi, A)$ . Note we have also used the fact that  $c(\pi, A) = p_{k(\pi, A)}^\pi(\cdot|A)$ .

Combining these two, we obtain

$$\begin{aligned} d_L(\pi, \pi') &= \hat{S}(\pi, \pi') + M \\ &= \hat{d}(\pi, \pi') + \hat{d}(\pi', \pi) + M \\ &> \hat{d}(\pi, c(\pi, A)) + \hat{d}(c(\pi, A), \pi) + M \\ &= d_L(\pi, c(\pi, A)) \end{aligned}$$

for all  $\pi' \in \Pi(\pi, A) \setminus \{c(\pi, A)\}$ . Hence, we have shown that  $c(\pi, A)$  is the unique solution to  $\min_{\pi' \in \Pi(\pi, A)} d_L(\pi, \pi')$ .

## 7.2 Proof of Proposition 6

The proof follows along lines similar to Claim 1 and Propositions 2 and 3. The satisfaction of the Independence axiom is needed only to guarantee that  $\succsim_A$  is well-defined whenever  $B(\succsim) \cap A \neq \emptyset$ .

**Proof of Part 1** : We show the first part of the proposition. Since  $|\Omega| \geq 3$ , there exist distinct  $x, y, z$  in  $\Omega$ . Now, let  $\succsim \in \mathcal{P}$  be such that  $B(\succsim) = \{x\}$ . Now, let  $A' = \{x, y\}$ ,  $B' = \{y, z\}$  and  $C' = \{z, x\}$ . Let  $\succsim' = c(\succsim, B')$ . Now by requirement, we must have  $B(\succsim') \subseteq \{y, z\}$ .

First note if  $B(\succsim') = \{x, y\}$ , then  $c((\succsim, B'), A')$  will equal the savage conditional preference of  $c(\succsim, B')$  based on  $A'$ . Hence,  $B(c((\succsim, B'), A')) = \{y\}$ . On the other hand, note that  $A'^c$  is  $\succsim$ -savage-null, hence  $A' \in \mathcal{K}(\succsim)$ . In this case, it follows that  $c(\succsim, A')$ , which is the savage conditional preference of  $\succsim$  based on  $A'$ , is unchanged and equals  $\succsim$ . But this means that  $B(c((\succsim, A'), B')) = B(c(\succsim, B')) = B(\succsim') = \{x, y\}$ . This violates strong path independence.

Now suppose that  $B(\succsim')$  is a singleton. W.l.o.g. assume that  $B(\succsim') = \{y\}$ . In this case, we

get  $B(c(\succsim, C'), A') = B(c(\succsim, A')) = \{y\}$ . However, it must be the case, by requirement, that  $B(c(\succsim, A'), C') \subseteq \{z, x\}$ . This is again a violation of strong path independence.

**Proof of Part 2** : It is straightforward to verify that every lexicographic updating rule satisfies Savage-consistency, AGM-consistency and weak path independence. We argue the "only if" direction. Suppose  $\langle c, \Pi \rangle$  satisfies Savage-consistency, AGM-consistency and weak path independence. Let  $\succsim \in \mathcal{P}$ . Since the updating rule is AGM-consistent, there is a complete transitive ordering  $\geq_{\succsim}$  corresponding to  $\succsim$ . Now, define the lexicographic preference profile  $\mathbf{P}^{\succsim}$  as

$$P_k^{\succsim} = c(\succsim, M_{\geq_{\succsim}}^k(\Omega)) \quad (10)$$

where  $M_{\geq_{\succsim}}^k(\Omega)$  is the set  $k$ -th highest elements in  $\Omega$  according to the ordering  $\geq_{\succsim}$ . We show that the family of lexicographic preference profiles  $\{\mathbf{P}^{\succsim}\}_{\succsim \in \mathcal{P}}$  indeed generates  $\langle c, \Pi \rangle$ .

Let  $A$  be non-empty. Let  $l$  be the smallest integer so that  $A \cap M_{\geq_{\succsim}}^l(\Omega) \neq \emptyset$ . Hence, since  $\langle c, \Pi \rangle$  is AGM-consistent, we have that  $B(c(\succsim, A)) = A \cap M_{\geq_{\succsim}}^l(\Omega)$ . This implies that  $\text{supp}\Pi(\succsim, A) \cap \text{supp}\Pi(\succsim, M_{\geq_{\succsim}}^l(\Omega)) \neq \emptyset$ . Now, note that since Savage-consistency is satisfied,  $c(c(\succsim, A), M_{\geq_{\succsim}}^l(\Omega))$  is the savage conditional preference of  $c(\succsim, A)$  based on event  $M_{\geq_{\succsim}}^l(\Omega)$ . But since  $B(c(\succsim, A)) \subseteq M_{\geq_{\succsim}}^l(\Omega)$ , we have that  $c(c(\succsim, A), M_{\geq_{\succsim}}^l(\Omega)) = c(\succsim, A)$ .

Now consider  $c(c(\succsim, M_{\geq_{\succsim}}^l(\Omega)), A)$ , which is simply the savage conditional preference of  $c(\succsim, M_{\geq_{\succsim}}^l(\Omega))$  based on  $A$ . However, from 10, this is equal to  $P_{l,A}^{\succsim}$ , which is the savage conditional preference of  $P_l^{\succsim}$  based on  $A$ . Finally, from weak path independence, it follows that  $c(\succsim, A) = P_{l,A}^{\succsim}$ .  $\square$